NAVAL POSTGRADUATE SCHOOL

MONTEREY, CALIFORNIA

THESIS

DECEPTION DETECTION PROCESS AND ACCURACY: AN EXAMINATION OF HOW U.S. MILITARY OFFICERS DETECT DECEPTION IN THE WORKPLACE

by

Kristofer A. Skidmore
Paul R. Ortiz

December 2014

Thesis Advisor: Lisa L. Massi Lindsey
Co-Advisor Juanita M. Rendon

Approved for public release; distribution is unlimited

THIS PAGE INTENTIONALLY LEFT BLANK

REPORT DOCUMENTATION PAGE			Form Approved OMB No. 0704–0188
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instruction, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188) Washington, DC 20503.			
1. AGENCY USE ONLY ***(Leave blank)***	**2. REPORT DATE** December 2014	**3. REPORT TYPE AND DATES COVERED** Master's Thesis	
4. TITLE AND SUBTITLE DECEPTION DETECTION PROCESS AND ACCURACY: AN EXAMINATION OF HOW U.S. MILITARY OFFICERS DETECT DECEPTION IN THE WORKPLACE		**5. FUNDING NUMBERS**	
6. AUTHOR(S) Kristofer A. Skidmore and Paul R. Ortiz			
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Naval Postgraduate School Monterey, CA 93943-5000		**8. PERFORMING ORGANIZATION REPORT NUMBER**	
9. SPONSORING /MONITORING AGENCY NAME(S) AND ADDRESS(ES) N/A		**10. SPONSORING/MONITORING AGENCY REPORT NUMBER**	
11. SUPPLEMENTARY NOTES The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government. IRB Protocol number NPS.2014.0051-IR-EP7-A.			
12a. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release; distribution is unlimited		**12b. DISTRIBUTION CODE** A	

13. ABSTRACT (maximum 200 words)

Research shows that humans are, on average, only slightly better-than-chance at deception detection. Meta-analysis conducted by Charles Bond and Bella DePaulo in their work "Accuracy of Deception Judgments" published by Personality and Social Psychology Review in 2006 yields an across-study average accuracy rate of 54%. Although prior research has failed to identify variables that have a large impact on accuracy, a recent set of studies focused on diagnostic utility (strategic questioning) leads us to expect substantial question effects producing levels of accuracy that differ substantially from chance. Recent research advocated for abandoning cue-based deception detection in favor of the idea of diagnostic utility. Specifically, this new line of research provides a basis for specifying the conditions under which questioning of honest and deceptive individuals yields levels of deception detection accuracy that depart substantially in both directions from the usual slightly-better-than-chance results that characterize past attempts. This thesis is a replication of these most recent diagnostic utility studies to determine if the methods are (1) generalizable to a new population and (2) useful in identifying specific questioning strategies relevant to Department of Defense and fraud investigation activities.

14. SUBJECT TERMS Deception Detection, Fraud Investigation, Diagnostic Utility, Questioning Methods, Strategic Questioning			**15. NUMBER OF PAGES** 67
			16. PRICE CODE
17. SECURITY CLASSIFICATION OF REPORT Unclassified	**18. SECURITY CLASSIFICATION OF THIS PAGE** Unclassified	**19. SECURITY CLASSIFICATION OF ABSTRACT** Unclassified	**20. LIMITATION OF ABSTRACT** UU

NSN 7540–01-280-5500 Standard Form 298 (Rev. 2–89)
Prescribed by ANSI Std. 239–18

THIS PAGE INTENTIONALLY LEFT BLANK

Approved for public release; distribution is unlimited

DECEPTION DETECTION PROCESS AND ACCURACY: AN EXAMINATION OF HOW U.S. MILITARY OFFICERS DETECT DECEPTION IN THE WORKPLACE

Kristofer A.Skidmore
Captain, United States Marine Corps
B.A., The Citadel, 2005

Paul R. Ortiz
Lieutenant, United States Navy
B.S., United States Naval Academy, 2007

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN MANAGEMENT

and

MASTER OF ARTS IN MANAGEMENT

from the

NAVAL POSTGRADUATE SCHOOL
December 2014

Authors: Kristofer A. Skidmore
Paul R. Ortiz

Approved by: Lisa L. Massi Lindsey
Thesis Advisor

Juanita M. Rendon
Co-Advisor

William R. Gates, Dean
Graduate School of Business and Public Policy

THIS PAGE INTENTIONALLY LEFT BLANK

ABSTRACT

Research shows that humans are, on average, only slightly better-than-chance at deception detection. Meta-analysis conducted by Charles Bond and Bella DePaulo in their work "Accuracy of Deception Judgments" published by *Personality and Social Psychology Review* in 2006 yields an across-study average accuracy rate of 54%. Although prior research has failed to identify variables that have a large impact on accuracy, a recent set of studies focused on diagnostic utility (strategic questioning) leads us to expect substantial question effects producing levels of accuracy that differ substantially from chance. Recent research advocated for abandoning cue-based deception detection in favor of the idea of diagnostic utility. Specifically, this new line of research provides a basis for specifying the conditions under which questioning of honest and deceptive individuals yields levels of deception detection accuracy that depart substantially in both directions from the usual slightly-better-than-chance results that characterize past attempts. This thesis is a replication of these most recent diagnostic utility studies to determine if the methods are (1) generalizable to a new population and (2) useful in identifying specific questioning strategies relevant to Department of Defense and fraud investigation activities.

THIS PAGE INTENTIONALLY LEFT BLANK

TABLE OF CONTENTS

LIST OF FIGURES

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF TABLES

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF ACRONYMS AND ABBREVIATIONS

ACFE	Association of Certified Fraud Examiners
DOD	Department of Defense
DIA	Defense Intelligence Agency
CIA	Central Intelligence Agency
CiC	Content in Context
FBI	Federal Bureau of Investigation
NPS	Naval Postgraduate School
PII	personally identifiable information
SUE	Strategic Use of Evidence

THIS PAGE INTENTIONALLY LEFT BLANK

ACKNOWLEDGMENTS

The authors wish to acknowledge the assistance provided by the many officers from all services who cooperated so generously in providing data. We are appreciative to Professor Marco DiRenzo, Ms. Naomi Overstreet, and Mr. H.G. Roser for their assistance in participant recruitment. It is with immense gratitude that we acknowledge the support and assistance of our advisors, Dr. Lisa L. Massi Lindsey and Dr. Juanita M. Rendon. Lastly, we wish to thank our families, who patiently tolerated long hours, piles of scattered documents, and collections of half-consumed cups of cold coffee. Without the unwavering support of all those named and unnamed this document would not have been possible. Thank you!

THIS PAGE INTENTIONALLY LEFT BLANK

I. INTRODUCTION

The body of work concerning interpersonal deception detection has, for decades, focused on the verbal and nonverbal cues of those being questioned. Regardless of the setting or sample, the methods of questioning subjects have not ventured beyond what the subject emits as cues. Over the course of the last decade, however, a new school of thought has emerged. Although there is an abundance of collected works on deception detection, this is the first thesis of its kind at the Naval Postgraduate School to study in detail the use of content- and context-based questioning within a military population. The issues discussed in this thesis have wide-ranging applicability both in the military and accounting/acquisition fraud environments. In evaluating how accurately military officers detect deception and what information is relied upon to make veracity judgments, the authors' overall research will be replicating previous studies conducted by Levine, Blair, and Clare (2014) and Park, Levine, McCornack, Morrison, and Ferrara (2002) on a previously unexamined population: domestic military officers. The successful completion of this thesis will not only add to the body of knowledge on the topic of deception detection but has the possibility for further replication in classified subsets such as defector interrogation and counterintelligence.

A. BACKGROUND/PROBLEM

The predominance of nonverbal cue-related research in the field of deception detection has led to stagnation in terms of accuracy improvement. Using the Levine et al. (2014) methodology of diagnostic utility, this thesis will show replicability and applicability within a military context. Furthermore, through applying the principles of Park et al.'s (2002) research, this thesis will examine how the population determines deception.

1. Deception Detection Accuracy

Until recently, research into deception detection has focused primarily upon interviewees' non-verbal cues. Decades of research on these non-verbal cues shows human beings are barely better than chance, garnering nothing more than a meta-analysis

accuracy rate of 53.46% (Bond & DePaulo, 2006). Most people believe in the fallibility of physical human nature, and that subtle nonverbal cues—the crossing of legs, the twitching of an eye, or the drumming of fingers—provide the critical clues to identifying when deception is occurring (Levine, Blair, & Clare, 2014). This can be seen on popular television programs, for example, when a player in a high-stakes poker match or an interrogated criminal suspect shows behavior that demonstrates leakage or a tell sign. Bond and DePaulo's (2006) most recent meta-analysis, however, looking at 208 deception detection accuracy studies, finds the opposite is true. When relying on nonverbal cues, people are not much better than chance (50%) in any given situation when attempting to determine if somebody is lying. In the literature, an accuracy ceiling of 65% exists (Levine et al. 2014).

The focal point of Study 1 is replicating research conducted by Levine et al. (2014), which indicates that more useful strategies can lead to better deception detection accuracy. Specifically, Levine et al. argue that what has been lacking for decades is attention paid to the specific question strategies used in interview situations. One important concept related to question strategies is diagnostic utility. Diagnostic utility is a scalable level upon which an individual uses information to form a correct conclusion. This scale ranges both positive and negative in that negative information would be viewed as deceptive. As noted by Levine et al. (2014) in their sixth experiment, which is being replicated in this research study (and which involves diagnostic utility), Levine and his colleagues were able to raise accuracy levels in excess of 70%.

2. How People Really Detect Lies

The prevailing deception detection literature and experimentation have been based on the immediacy of non-verbal cues exhibited by an individual and interpreted by a recipient. Though this concept retains validity, it is inherently flawed as it fails to address the additional information a recipient possesses when he/she makes a determination of authenticity. As examined by Park, Levine, McCornack, Morrison, and Ferrara (2002), most individuals in lie determination settings rely upon information gathered prior to the detection of the lie. This elongated timeline includes "information

sources such as information from third parties, the consistency of statements with prior knowledge, the consistency of messages with physical evidence, and confessions," which when taken into account with non-verbal cues build an individual's truth/lie judgment (Park et al., 2002, p. 144). It is this combined effect of non-verbal and subjective historical evidence that forms the basis of Study 2.

3. The Current Studies

Study 1 replicates the sixth study in Levine et al.'s (2014) program of research using U.S. military officers to determine whether a replication with a different population would yield similar results and, thus, determine the generalizability of their findings. Study 2 replicates the Park et al. (2002) research, which that involved the applicability of historical subjective information in the detection of lies, to see if such information varies when applied to a different population. The results of both studies will help to better shape the field of interrogation, defector questioning, and fraud detection in U.S. military contexts.

B. PURPOSE

The purpose of this research is to conduct the experimental replication and documentation of further improvement in deception detection accuracy over that of previous findings. The nature of this topic dictates the use of both qualitative and quantitative opportunities to further expand known knowledge of deception detection. Quantitatively, the focus will be on strategic questioning methods and resulting deception detection accuracy. Qualitatively, this research looks at when, how, and with what information deception is actually detected in the workplace.

1. Study 1: Content and Context Questioning Effects on Accuracy

Unlike the bulk of the previously conducted research, Study 1 aims to remove the focus on non-verbal cues. Question effects detail how the questioning of an individual may potentially impact the outcome of deception detection. It is specifically the manipulation of such question effects through the use of applied content and contextual questioning that the authors believe can and will lead to an increase in diagnostic utility.

Diagnostic utility is the overarching conceptual idea that information has varying amounts of utility, both positive and negative. It is the application of diagnostically useful statements that the authors measure as a means of deception detection.

2. Study 2: How People Really Detect Lies

The purpose of Study 2 is to examine what additional information the selected population rely upon in their truth/lie determinations and the associated time horizon with regards to deception realization. This study differs from traditional works in that, rather than focus on the questions or questioning method, it instead places emphasis on the individual detectors' individual backgrounds and historical subjectivity.

C. OBJECTIVES

As our primary objective, this research study will provide a thorough and current review of the issues regarding deception detection and its applicable utility within the Department of Defense and other governmental agencies. Study 1 will seek to determine the role of specific question strategies and expertise (previous interview or interrogation training/experience) on deception detection accuracy by replicating Levine et al.'s (2014) research on diagnostic utility. Study 2 will seek to determine what subjective historical methods individuals use in deception detection by replicating Park et al.'s (2002) work on what information people use other than the verbal and non-verbal behaviors of the liar when determining deception.

D. RESEARCH QUESTIONS

This thesis and associated studies are, in essence and design, a replication of Levine et al.'s (2014) and Park et al.'s (2002) studies to determine if both researchers' results are generalizable beyond college students and law enforcement and useful in identifying specific questioning strategies relevant to DOD activities to include defector questioning and fraud examination. Specifically, this thesis aims to answer the following two research questions:

- Utilizing the diagnostic utility methods of content and contextual questioning presented in the Levine et al. (2014) study, are Department of Defense officers

able to distinguish deception with greater accuracy than the previous meta-analysis mean of 54%, and if so to what degree?

- Utilizing the subjective historical reliance method outlined in Park et al. (2002), what types of information do U.S. military officers report using when detecting lies in the workplace?

E. SCOPE

This research project examines the abilities of U.S. military officers (N = 84) to detect deception when presented with high-stakes interviews where some interviewees lied and others told the truth. Specifically, Study 1 examines the role of specific questioning strategies and the expertise of the participants (expert vs. inexpert) in accuracy levels. Study 2 examines what, if any, additional subjective information participants use in determining deception and over what associated timeline did veracity judgments occur. This thesis merely looks to replicate both the Levine et al. (2014) and Park et al. (2002) research and determine whether either study is generalizable to a different population, namely U.S. military officers. Also, the results of these studies might shed light on appropriate high-stakes deception settings such as defector questioning and detection of fraud in DOD procurement and acquisition settings.

F. ASSUMPTIONS AND LIMITATIONS

(1) Assumptions

As explained in both the Levine (2014) and Bond and DePaulo (2006) works, the primary assumptions of this research include that recent findings are not flukes and are instead related to changes in the research and empirical findings of older limited works. Additionally, this research study assumes domestic military officers are a distinct and different population from what has previously been studied. This uniqueness is a result of training received by military personnel, the inherent cultural emphasis on truthfulness, and risks associated with high-stakes deception. Lastly, another assumption is that Study 2 will replicate the work of Park et al. (2002), given that Lindsey et al. (2012) is the only other existing replication of such results using a workforce sample.

(2) Limitations

Previous deception studies assume participants will display a truth bias, in that human beings have shown over time a belief in the innate honesty of others (Park et al. 2002). This explains the feelings of betrayal often exhibited by those who have experienced lies. For the purpose of this research, the authors assume the participants will operate with a truth bias. However, truth bias is not manipulated or measured to determine if that is a replicable finding in the specific population studied. This thesis will not address any psychological factors such as mood, temperament, or any other conditions that might affect respondents' participation. No financial constraints limited the authors, as data collection included only volunteer participation.

G. ORGANIZATION

This thesis is organized into two separate, but mutually reinforcing, studies. Chapters II–V comprises Study 1. Study 1 is the replication of Levine et al.'s (2014) study of diagnostic utility—or questioning effects—on the role of expertise on detection deception accuracy. Specifically, Chapter II is the literature review that introduces the history of deception detection research and related literature. Chapter III provides the Study 1 methodology, including information on participants, materials, and the procedure. Chapter IV presents the statistical results of Study 1, and Chapter V expands on those results through detailed discussion of the findings and their limitations.

Chapters VI–IX comprises Study 2. Study 2 is the replication of Park et al.'s (2002) study of how people really detect lies in their everyday lives. Chapter VI is the literature review that introduces the premise of Park et al.'s (2002) research on how people actually detect lies in their interactions. Chapter VII provides the Study 2 methodology, including information on participants and the procedure. Chapter VIII presents the results of Study 2, and Chapter IX expands on those results through detailed discussion of the findings and their limitations.

Lastly, in Chapter X, the overall summary, conclusions, and recommendations for further deception detection research will be presented.

II. STUDY 1: BACKGROUND HISTORY/LITERATURE REVIEW

Few things are as fundamentally human as the quest to accurately ascertain the veracity of one's intentions. The earliest codified documents are rife with tales of, and punishments for, lying; be it from Indian Sanskrit Vedas or the Greek physician Erasistratus, humanity's ordeal with the nature of truth remains a constant (Trovillo, 1939).

A. HISTORICAL BACKGROUND OF DECEPTION DETECTION

The historical underpinnings of deception detection begin with the initial reliance on what Trovillo (1939) notes as superstition and the concept of the Ordeal.

> It is significant that, with few exceptions, the historical accounts of deception-detecting from the days of Christ, through the Middle Ages, are the history of the Ordeal. Superstition so swayed the minds of people that it was the rule for them to ask for the Ordeal to prove their innocence. The accuser was not looking, evidently, for suspicious clues in the face or actions of the individual, for apparently the psychology of deceit did not exist. Even the religions of Europe, as late as the 16th Century, taught that proof of innocence or guilt would be furnished from on High in a variety of mystical modes. People did not consider that proof lay within or on the surface of the suspect himself. (p. 850)

The Ordeal method of deception detection is easily understood when examined in the context of the early witch trials and the Inquisition of the 14th and 15th centuries where individuals were forced to participate in torturous tests where the outcome of the event was the prime determinate of the presence of deception.

As religious fervor subsided and use of the scientific method began to flourish, so too did the study of emotional states and their bearing on deception detection (Trovillo, 1939). One of the earliest such works was conducted by Mosso, an Italian physiologist. Mosso's work focused on the effect of fear with regard to deception, specifically the fear one experiences in being detected. Mosso's measurement of blood flow as it circulates and pools in the body led to the first crude attempts to measure the physiological effects of deception (Trovillo, 1939).

What these two examples illustrate is the ever shifting framework through which deception detection has been viewed. Recent, cutting-edge research indicates that diagnostic utility—or the degree to which information is useful as prompted through strategic questioning—is key to understanding how humans detect deception. Prior theories, however, hinge upon the psychological and physiological states of the person lying, and the resulting nonverbal cues that could be "read" to detect deception. The recent work on diagnostic utility questions the usefulness of these nonverbal cues that might or might not yield diagnostically useful information about whether somebody is lying. Such reliance on nonverbal cues has yielded detection accuracy rates that are not much different than chance (Bond & DePaulo, 2006). As such, the concept of diagnostic utility presented in this thesis is based on Levine et al.'s (2014) work and differs from the classical interpretation of diagnostic utility in that it includes the use of Park et al.'s (2002) additional reliance on subjective historical information. Specifically, Levine et al. (2014) indicate that diagnostic utility, in relation to deception detection, is the consideration of contextual message content including plausibility, correspondence with known facts and the correlation of the two, and these are what influence one's true ability to detect deception.

B. REVIEW OF PREVIOUS DECEPTION DETECTION ACCURACY STUDIES

Upon review of the last four decades of deception detection research, a common theme emerges across all studies: a set of individuals is recruited as message initiators (liars or truth tellers), a separate group acts as the sample and is tasked to determine the veracity of the initiator's message with accuracy being calculated as the proportion of correct judgments made by the sample and based on non-verbal low-risk deception settings (Park et al. 2002). Furthermore, the meta-data accuracy rate presented by these studies consistently falls around 57% (Kraut, 1980) and always between 45% and 70% (e.g., Kalbfleish, 1994; Miller & Stiff, 1993; Vrij, 2000). Over the past 40 years, and 208 studies as noted by Bond and DePaulo (2006), this belief has been near collectively held in the literature (e.g., Burgoon, Buller, Ebesu, & Rockwell, 1994; DePaulo, Kirkendol,

Tang, & O'Brien, 1988; Feeley, deTurck, & Young, 1995; Kalbfleisch, 1994; Millar & Millar, 1995; Stiff & Miller, 1986; Vrij, 1994).

What deviated from this trend was the work by Levine et al. (2014) that focused on the previously described intrinsic value of diagnostic utility. Two complementary approaches to deception detection involving diagnostically useful information comprise the current pertinent body of knowledge concerning the subject. The Strategic Use of Evidence (SUE) method hinges on an interviewer possessing a form of useful evidence pertinent to the line of questioning without the interviewee knowing, in hopes the subject will inadvertently make false statements as compared to the evidence (Clemens, Granhag, & Strömwall, 2013). Further analysis regarding the way in which the evidence is presented during the interview leads to what has been described earlier as the Content in Context (CiC) technique. Both methods set an arbitrary baseline with presumably useful obtained, or obtainable, knowledge that allows the interviewer to measure the variable feedback (Levine et al., 2014). Providing useful knowledge and background specifics are critical to both methods and at the heart of this current study.

C. STUDY DESIGN

This study was a 3x2 mixed design with the three sets of questions presented in Table 3 as a repeated factor, the two levels of expertise (expert versus inexpert) as an independent groups variable, and detection accuracy as the dependent variable. Expertise was operationally defined using participants' answers to the following question: Have you ever conducted interviewing or interrogation as a regular part of your job? Participants who answered "No" were coded as inexpert; those who answered "Yes" were coded as experts. An additional question asked was, "Have you ever received formal interviewing or interrogation training?" Only 20% of participants had received such training; therefore, the sample size was too small to make meaningful comparisons in the current study using this question as a proxy for expertise. Participant answers (their truth/lie judgments) were scored for accuracy by adding the number of correct judgments and dividing by the total number of judgments.

Each participant watched and rated 12 videotaped interviews of different students denying cheating. The 12 interviews used in the current study were the same interviews that were used in Levine et al.'s sixth study. For each of the question sets, two lying cheaters' interview segments were shown. For each deceptive interview, a corresponding honest interview was selected by matching on sex, race, and approximate physical appearance. Thus, there were two honest non-cheaters and two lying cheaters, all of whom denied cheating, interviewed with each of the three question sets.

D. RATIONALE FOR STUDYING DECEPTION DETECTION ACCURACY IN THE U.S. MILITARY

Initial thoughts on deception detection in the military predominately involve intelligence and counter-intelligence specialty fields. As previously noted, however, the humanistic quality of deception means detection practices are relevant in all settings of interpersonal interaction. As a government entity, the military must always maintain ethical practices and morally sound principles, which in turn facilitate a strong truth bias. It is very difficult for individuals interacting within the military system, where it is reasonably assumed that all participants are honest brokers, to then exit the system and work with individuals whose motivations are self-serving and not bound by the same politico-social contract. Thus, deception is easily found in various forms in different specialties. Not all specialties or circumstances require deception detection training on the level of enemy combatant interrogation. Deception detection, however, is a useful tool in areas such as contracting fraud, military law enforcement, and varying degrees of leadership where high stakes deception can occur and where a healthy skepticism and base of training may aide mission accomplishment.

Buller's (1996) research on Interpersonal Deception Theory for the U.S. Army Research Institute serves as the primary academic linkage between this work's analysis of deception detection theory and the military, and his four-year examination inadvertently parallels the same issues as the research questions posited in this work. Buller's (1996) focus on issues related specifically to intelligence gathering personnel will be addressed further in Chapter X, which explores defector deception detection and the application or

influence of psycho-cultural lenses on strategic posturing/positioning as a focus of further research.

E. RESEARCH QUESTIONS

Study 1 research is based upon the qualitative replication of the Levine et al. (2014) methodology. Utilizing the diagnostic utility methods of content and contextual questioning presented in the Levine et al. (2014) study, are Department of Defense officers able to distinguish deception with greater accuracy than the previous meta-analysis mean of 54% and if so to what degree? If the study holds true to the findings of Levine et al. (2014), then the authors should find a mean accuracy in excess of 54% and improvement most likely greater than 70%, with a corresponding minimal effect for expert judgments. Should the replication prove false, further research will be required to examine which portion of the study does not hold true in the chosen sample and to what amount modification will need to occur. The next chapter fully details the methodology used in Study 1.

THIS PAGE INTENTIONALLY LEFT BLANK

III. STUDY 1: METHODOLOGY

In order to replicate the Levine et al. (2014) experiment 6 in an appropriately clinical method, the authors first sought approval of the NPS Institutional Review Board for the Protection of Human Subjects (IRB). The authors completed the mandated IRB ethically based training and all additional reviews, and subsequently received approval to initiate research. All research occurred on campus in specially designated rooms designed to best replicate the atmosphere of the initial study. Daily findings and survey materials were maintained under secure conditions, and no leakage of personally identifiable information (PII) or demographics occurred.

A. PARTICIPANTS

The participants in this study were 84 U.S. military officers studying in resident programs at the Naval Postgraduate School. Tables 1 and 2 summarize participants' rank and branch of service, respectively. Participants' ages ranged from 22–51 (M = 33.74, SD = 4.86) and years of military service ranged from 1–27 (M = 12.21, SD = 5.31). Also, 76.2% were male, 77.4% identified themselves as Caucasian/White, 8.3% were African American/Black, 4.8% were Asian/Pacific Islander, 8.3% were Hispanic/Latino(a), and 1.2% indicated they were multiple races/ethnicities. In terms of expertise, 22.6% indicated they had received formal interviewing or interrogation training, and 40.5% reported that they had conducted interviewing or interrogation as a regular part of their jobs. All participants volunteered their participation and none received any form of compensation for their involvement in this research. Participation was anonymous with the only demographic information being gender, age, race, rank, years of military service, military branch, and level/use of interrogation training.

Table 1. Participant Military Rank

RANKS	FREQUENCY	PERCENT
CW04 (Chief Warrant Officer 4)	1	1.2%
O1 (Second Lieutenant/Ensign)	5	6%
O2 (First Lieutenant/Lieutenant Junior Grade)	3	3.6%
O3 (Captain/Lieutenant)	55	65.5%
O4 (Major/Lieutenant Commander)	18	21.4%
O5 (Lieutenant Colonel/Commander)	2	2.4%

Table 2. Participant Branch of Service

BRANCH OF SERVICE	FREQUENCY	PERCENT
Army	11	13.1%
Navy	41	48.8%
Air Force	10	11.9%
Marines	21	25%
Coast Guard	1	1.2%

B. MATERIALS

Twelve videotaped interview segments, which were all used in the Levine et al. (2014) examination of diagnostic utility and expertise in their sixth study, were used as the experimental stimuli in this current study. Specifically, each segment showed a university student being interviewed about cheating in a research-related trivia game. All students in the 12 segments denied cheating, but half (6) actually cheated. Thus, the 12 video segments resulted in six honest students telling the truth (they denied cheating and they, in fact, did not cheat) and six dishonest students lying (they cheated but denied it in

the interview). See Levine et al. (2014, especially pp. 270–272) for a detailed description of how the experimental videotapes were created. The 12 interview segments represented three different diagnostic utility questioning strategies (see Table 3).

Table 3. Questions Asked in the Three Sets of Cheating Tapes (from Levine et al., 2014, p. 268)

First Question Set
1. Did you find the trivia questions difficult? 2. Was teamwork much of a factor? How so, please explain. 3. In looking at your score, you did better than other groups. Does this surprise you? 4. How would you explain your success? 5. Did cheating occur when I left the room? 6. Why should I believe you?
Second Question Set
1. How much experience have you had working in teams? 2. Have you played trivial pursuit games before? 3. How often do you play? [follow-up] Why do you/don't you play them much? 4. In the current game, did you find the questions difficult? [follow-up] If you were going to scale it on a scale of 1 to 10, what would you say they were in terms of difficulty? (1 being easy and 10 being difficult) 5. Was teamwork much of a factor in your performance? 6. What would you say if I told you that you did better than other groups? 7. How would you explain your performance? 8. Did you cheat when the experimenter left the room? 9. Are you telling me the truth? 10. What would your partner say if I asked them the same question (question 9)?
Third Question Set
1. Tell me, in as much detail as you can, what happened during the trivia game? 2. How well did you do on the trivia game? 3. Which questions did you and your partner get right? 4. For the answers you got right, explain how you knew the right answer? 5. In detail, what happened when the experimenter left the room? 6. Did any cheating occur? 7. When I interview your partner, what will he/she say about cheating? 8. Did you and your partner discuss cheating? 9. If someone did cheat, what should happen to him/her?

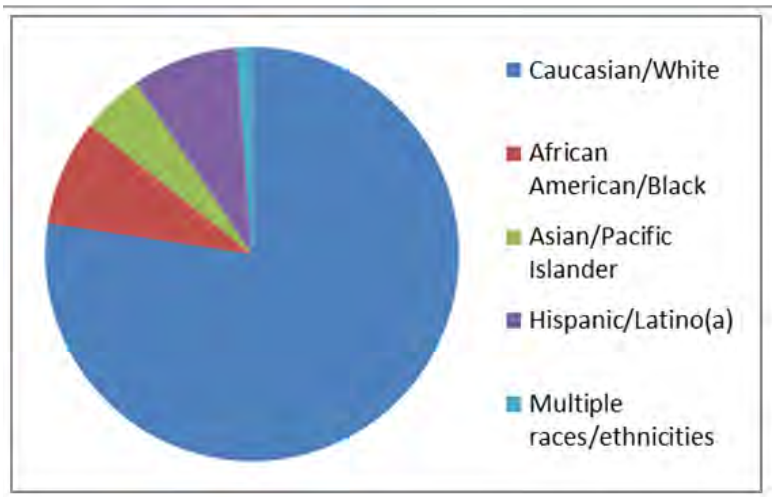
C. PROCEDURE

Participants entered a lab setting, at which point they read and completed a consent document. After consenting to participate, each respondent watched a series of 12 videotaped interview segments lasting approximately two minutes apiece. After each segment, the participant paused the video for as long as necessary to answer survey questions to make a truth-lie judgment regarding the individual in the video (see Appendix A for full questionnaire for Studies 1 and 2). Participants also answered a series of demographic questions.

D. DATA ANALYSIS APPROACH

Data were analyzed using mixed-model Multivariate Analysis of Variance, where question strategy was the repeated factor and expertise was the between factor, both comprising the independent variable and with accuracy as the dependent variable. Effect sizes are also reported.

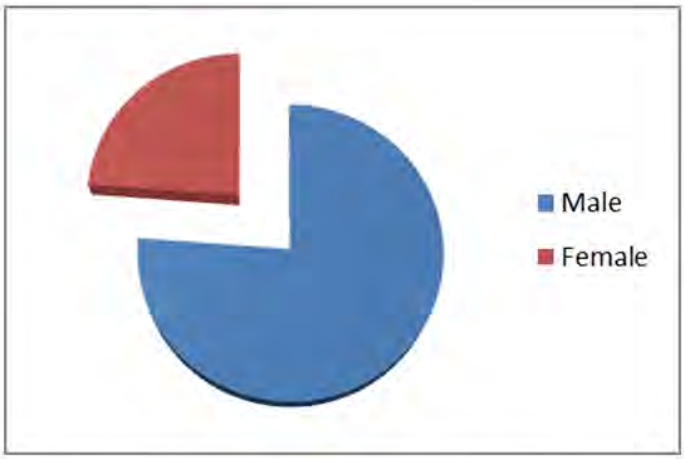
While the majority of the sample identified as Caucasian/White as shown in Figure 1, this is in keeping with the appropriate population percentages as published by The Department of Defense and based on 2010 Active Duty United States Military member personnel records (CNA Corporation, 2012). Additionally, the sample is in congruence with Naval Postgraduate School ethnicity demographics as published in the 2013 NPS Factbook (Naval Postgraduate School, 2013). The small sample size for races other than Caucasian/White does not allow for statistically significant analysis to be conducted.



Race/Ethnicity		
	Frequency	Percent
Caucasian/White	65	77.4
African American/Black	7	8.3
Asian/Pacific Islander	4	4.8
Hispanic/Latino(a)	7	8.3
Multiple races/ethnicities	1	1.2
Total	84	100

Figure 1. Race/Ethnicity

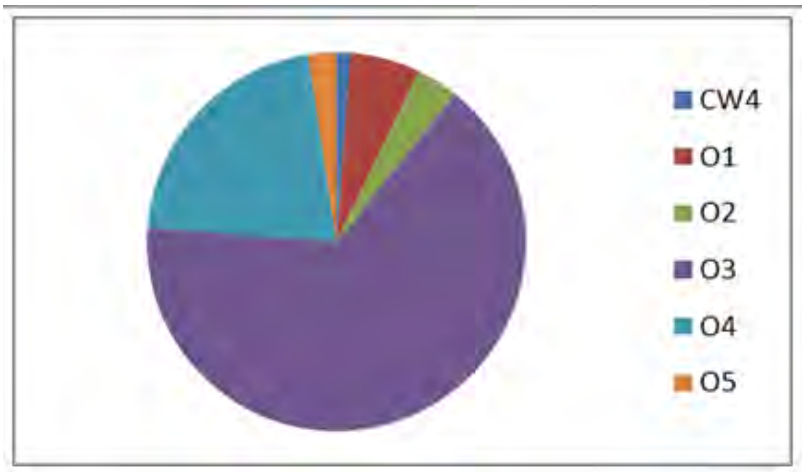
As with race/ethnicity, the gender disposition as shown in Figure 2 is in keeping with the aforementioned 2010 Department of Defense Demographics Report. The sample size for this study, however, included a slight increase in female participation, percentile specific (CNA Corporation, 2012). This increase also holds true when examining the NPS population (Naval Postgraduate School, 2013).



Gender		
	Frequency	Percent
Male	64	76.2
Female	20	23.8
Total	84	100

Figure 2. Gender

As shown in Figure 3, the sample was composed of a diverse range of military ranks centered predominately on mid-career officers, O-3 and O-4, which is in keeping with the overall demographics of the school (Naval Postgraduate School, 2013). A mean time in service of 12.21 years and a standard deviation of 5.3 years coupled with the mean age of 33.74 years and a standard deviation of 4.8 years lend credibility to the belief that the sample is sufficiently experienced but may have an impact regarding the biases of the individuals.

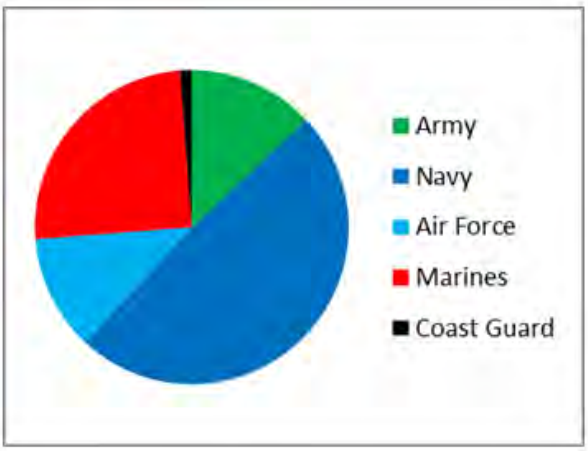


Rank		
	Frequency	Percent
CWO4	1	1.2
O1	5	6
O2	3	3.6
O3	55	65.5
O4	18	21.4
O5	2	2.4
Total	84	100

	N	Minimum	Maximum	Mean	Std. Deviation
Age	84	22	51	33.74	4.857
YearsOfService	84	1	27	12.21	5.314

Figure 3. Rank

As shown in Figure 4, the sample contained a diverse mixture of military branches loosely following the represented population. There was a slightly larger sampling of Marine Corps officers as compared to Army. The population typically holds the Army officer contingent slightly larger than that of the Marine Corps (Naval Postgraduate School, 2013).



Branch		
	Frequency	Percent
Army	11	13.1
Navy	41	48.8
Air Force	10	11.9
Marines	21	25
Coast Guard	1	1.2
Total	84	100

Figure 4. Branch

Using the mixed-model Multivariate Analysis of Variance with question strategy and expertise as the independent variables and accuracy as the dependent variable proved to be a logical method for data analysis, which lends itself to robust replication and expansion in further research. The next chapter discusses the results of Study 1.

IV. STUDY 1: RESULTS

The data were analyzed with a 3x2 mixed Analysis of Variance with the three question sets as a repeated factor, the two levels of expertise (expert vs. inexpert) as an independent groups variable, and detection accuracy (percent correct) as the dependent variable.

Consistent with Levine et al. (2014), this study replicated the strong main effect for questioning strategy, $F(2, 164) = 61.56$, $p < .001$, $\eta^2 = .43$. Although Levine et al. found a statistically significant main effect for expertise and a statistically significant question type x expertise interaction, this research study did not (one should note that the effect sizes for both of Levine et al.'s findings were trivial with both $\eta^2 = .01$). Specifically, neither the main effect for expertise in this research study, $F(1, 82) = 1.79$, $p = 0.18$, nor the question by expertise interaction were statistically significant, $F(2, 164) = 0.05$, $p = 0.95$.

The cell means are presented in Table 4. Across experts and inexperts, accuracy was 37.5% (95% CI = ± 2.8%), 67.8% (95% CI = ± 2.6%), and 78.4% (95% CI = ± 2.5%) for question sets 1, 2, and 3, respectively. Furthermore, accuracy means in all six cells were significantly different from both 50–50 chance and the 54% meta-analysis mean at $p < .01$.

Table 4. Mean Accuracy (and Standard Deviations) by Condition, Study 1

Question Set	**Set One**	**Set Two**	**Set Three**
Expert	36.0% (23.2%)	66.2% (22.9%)	75.7% (22.6%)
Inexpert	39.0% (26.8%)	69.5% (23.3%)	81.0% (22.3%)

Participant veracity precision is not dependent upon expertise and increases both with the addition of diagnostically useful information and at a similar rate to that expressed in Levine et al. (2014). The following chapter provides discussion, findings, limitations, and recommendations based on analysis.

THIS PAGE INTENTIONALLY LEFT BLANK

V. STUDY 1: DISCUSSION, FINDINGS, LIMITATIONS, AND RECOMMENDATIONS BASED ON ANALYSIS

The results, as hypothesized, replicated the previous findings illustrated in Levine et al.'s (2014) research. The importance of this is that the findings exhibited an increase in deception detection irrespective of expertise; both expert and inexpert categories saw a marked increase in accuracy as previously shown in Table 4.

A. DISCUSSION AND FINDINGS

In light of the literature review, typical physical and nonverbal variables saw a replicable meta-data analysis mean of 54% (Bond & Depaulo, 2006), while the authors' replicable method of manipulating question effects through the use of applied content and contextual questioning raised expert participant accuracy to 75.7% and inexpert participant accuracy to 81.0%. The accuracy increase of this replication using the previously defined sample of domestic military officers actually saw a notable increase in accuracy over the Levine et al. (2014) results, especially with regards to the inexpert category. The Levine et al. (2014) mean accuracy results are provided in Table 5 for comparison.

Table 5. Mean Accuracy (and Standard Deviation) by Condition (from Levine et al., 2014)

Question Set	Set One	Set Two	Set Three
Experts	29.2% (22.9%)	64.3% (24.1%)	72.4% (22.2%)
Student (Inexpert)	39.2% (24.1%)	66.6% (24.9%)	72.9% (22.7%)

Additionally, the Levine et al. (2014) finding of "negative utility answers seem to hinder experienced judges more than positive utility answers assist experienced judges" holds true in this study's replication and is an interesting issue to address in follow-on research and analysis.

B. LIMITATIONS

Though the sample size (N = 84) could be considered small, it was directly representative of both the Naval Postgraduate School domestic officer population (Naval Postgraduate School, 2013) and comparatively demographically as diverse as the Department of Defense (CNA Corporation, 2012) and yet, despite the sample size, elicited a very strong positive accuracy effect when incorporating the new variables. This proves the usefulness and robust nature of the strategic questioning method.

Additional limitations include the restriction of the sample to that of domestic officers and the single contextual nature of the lies examined in the experiment. The nature of deception in the military and financial fraud environments are inherently more high stakes than that of academic cheating.

C. RECOMMENDATIONS BASED ON ANALYSIS

The value of this research is in applying the methodology to future deception detection training for inexpert domestic military officers. The use of contextual and content based questioning was most beneficial to the inexpert sample meaning initial training of deception detectors should be focused primarily on this style. Follow-on or remedial training for expert detectors should incorporate elements of diagnostic utility. However, trainers should not expect the same return as when used with inexpert subjects. The next chapter will provide the background and literature review for Study 2.

VI. STUDY 2: BACKGROUND HISTORY/LITERATURE REVIEW

The focus of Study 2 is the qualitative factor of deception detection vice the quantitative issue of accuracy as in Study 1. Study 2 seeks to determine what factors, in addition to non-verbal cues and leakage, individuals use when making a truth/lie judgment. As Park et al. (2002) argue that there are four false assumptions that previous research has relied on in believing the primacy of verbal and non-verbal behaviors of the messenger—"questions researchers have asked, the research designs used to study deception detection, the directions the literature has taken, and the theories used to predict and explain the results" (p. 147)—Study 2 directly addresses the issues of the first and second assumptions by broadening the field of questions asked in relation to the underpinning reasons for the truth/lie judgment and by modifying the study to attempt replication of the Park et al. (2002) findings.

The only other study of workplace deception detection, Lindsey, Dunbar, and Russell (2011), specifically examined the relational aspect of the power dynamic to that of deception detection. The study consisted of a sample of 214 employed individuals of which 55% were in a management/supervisor status, of which (N = 96) stated they had engaged in deception in the workplace (Lindsey et al., 2011). Remarkably, "no lies were uncovered through the interpretation of nonverbal cues, rather they were discovered after the fact through evidence or confessions" (Lindsey et al., 2011, p. 74). This effect of after-the-fact evidence and complete disregard for nonverbal variables coupled with the workplace power dynamic lends great weight to attempting a replication of the study within the highly charged culture of the military, where both power and stakes are greatly increased.

Further support of investigating the qualitative rationale for veracity judgment is provided by Park et al., who argue in *How People Really Detect Lies* (2002) that the majority of previous studies relied too heavily on the nonverbal cues of interviewees in addition to factors such as:

(1) Sources and judges in deception detection experiments were most often unacquainted and that detection accuracy might be higher if judges had relational or idiosyncratic knowledge of the message source.

(2) Participants had only rarely been allowed to interact face-to-face.

(3) Lies are often sanctioned (encouraged to a degree) by the researcher. This argument holds that liars telling sanctioned lies should be less aroused than those telling unsanctioned ones. Consequently, unsanctioned lies should be more easily detected than sanctioned lies, and detection accuracy might be better if more researchers studied unsanctioned lies.

(4) Predominance of testing under conditions of everyday vice high-stakes lies.

Awareness of these limiting assumptions must be noted when conducting deception research. The risk associated with the factors listed above can be mitigated through the use of content- and context-based questioning regimens.

B. RESEARCH QUESTION

Utilizing the Park et al. (2002) method, what types of information do U.S. military officers report using when detecting lies in in the workplace? If the study holds true to the original as set forth by Park et al. (2002), the authors expect to find that the vast majority of individuals report using subjective historical information or after-the-fact data rather than non-verbal cues to detect deception. The following chapter discusses the methodology used for Study 2.

VII. STUDY 2: METHODOLOGY

Replication of Park et al.'s (2002) study occurred in an appropriately clinical manner, beginning with the authors' approval of the NPS Institutional Review Board for the Protection of Human Subjects (IRB). The authors completed the mandated IRB ethically based training and all additional reviews, and subsequently received approval to initiate research. All research occurred on campus in specially designated rooms designed to best replicate the atmosphere of the initial study. Daily findings and survey materials were maintained under secure conditions, and no leakage of personally identifiable information (PII) or demographics occurred. Upon completion of the study, the questionnaires were divided, and all qualitative information was independently coded by two coders.

A. PARTICIPANTS

Participation in Study 2 was fully voluntary and anonymous, and comprised the identical sample used in Study 1: U.S. military officers (N = 84) serving as students at the Naval Postgraduate School. Tables 1 and 2 and Figures 1, 2, 3, and 4 discussed in Chapter III for Study 1 provide the full details and breakdown of the participant demographics.

B. PROCEDURE

After completing the truth-lie judgments in Study 1, respondents were asked a series of open-ended questions following Park et al.'s (2002) protocol (see Appendix: Consent Form and Questionnaire). Specifically, participants were told to recall a recent work-related situation in which they discovered that someone lied to them, and they were asked to remember as much as they could about what happened.

Participants were asked to write a detailed description of the event:

(1) Recall as much as you can about the situation in which the person originally lied to you. In as much detail as possible, describe the event

where you were lied to: Where did it happen? What was the lie about? If you can, be sure to write down the exact thing that the person said to you.

Next, respondents were asked a series of questions related to some of the details surrounding the situation:

(2) How long ago did this event (the lie) originally take place?

(3) What was/is the relationship between you and the person who lied to you?

(4) Now, think about how you found out that the person lied to you. Describe in as much detail as you can the events surrounding your discovery of the lie: how exactly did you find out that the person lied to you?

2. Coding of Qualitative Data

All questionnaires were collected from participants upon conclusion of their session and independently coded by two coders. The coding scheme was created by the authors based on N = 84 data collected from willing United States military officers serving in resident student capacity at the Naval Postgraduate School. Participation was fully voluntary and anonymous and comprised the identical sample used in Study 1. Upon coding completion, the inter-coder reliability (Kappa) was calculated and any discrepancies were resolved via discussion between coders and the primary thesis advisor. This data is further reported in the results section. The qualitative codebook comprises the following questions and their associated Kappa. The question addressing "how long ago the lie was originally told" was coded using months as the unit of measure with a resulting Kappa of .99. The question of linking "relationship" included none specified, superior/immediate boss, superior/above immediate boss, subordinate, child, spouse, immediate family member (brother, sister, mom, dad), peer/friend, teacher/caregiver, senior in rank (but no command relationship), and other. Relationship coding resulted in a Kappa of .91. The "discover method" coding options included none listed, 3rd party information, physical information, solicited direct confession, unsolicited direct confession, at-the-time verbal and/or nonverbal behavior, inconsistencies with prior knowledge, combination of two or more, and other. The "discovery method" Kappa

was .84. The question addressing the lapse of time between when the lie was told and when the subject discovered the lie was coded as no answer provided, immediate detection, less than one hour, less than one day, less than one week, less than one month, less than one year, and more than one year. The Kappa for time lapse was .72. The second coder transcribed the listing of discovery method examples as presented in Chapter IX. The following chapter will discuss the results of Study 2.

THIS PAGE INTENTIONALLY LEFT BLANK

VIII. STUDY 2: RESULTS

Study 2 respondents were asked to recall a work-related situation in which an individual lied to the participant; data was not collected or analyzed for 36 (43.4%) respondents who choose to recuse themselves of the question. Additionally, one respondent noted deception by their child, as the current study is interested in workplace related deception this interfamilial deception was eliminated from further analysis. Therefore, the sample for the relational research question is N=84 while the rest are N=83 to compensate for the removal of the non-workplace deception respondent.

For those who chose to answer, 48 (57%) of the initial sample, the most common discovery methods were 3rd Party Information (16.9%) and Physical Information (14.5%). The least common valid discovery method, excluding one answer (1.2%) that was not categorized, was At-The-Time Verbal/Nonverbal Behavior (0.0%). Table 6 delineates these findings while Table 7 presents examples of the discovery methods.

Table 6. Frequencies of Recalled Workplace Lie Discovery Methods, Study 2

Discovery Method	*f*	%
None Listed	36	43.4%
Third Party Information	14	16.9%
Physical Information	12	14.5%
Solicited Direct Confession	2	2.4%
Unsolicited Direct Confession	4	4.8%
Verbal/Nonverbal Behavior	0	0.0%
Inconsistencies with Knowledge	5	6.0%
Combination	9	10.8%
Other	1	1.2%
Total	83	100.0%

Table 7. Examples of Discovery Method Categories

DISCOVERY METHOD	RESPONSE EXAMPLE
Third Party Information	A subordinate naval officer lied about adhering to the stipulations of temporary assigned duty orders, a peer of the deceiver informed the respondent of the lie.
Physical Information	An individual stated they had moved a class however, a computer system query by the respondent proved this to be false.
Solicited Direct Confession	In response to an inquiry as to why an enlisted sailor could not work a New Year's Day shift the sailor confessed during non-judicial punishment that he had lied about being sick during the shift and instead was too hung-over to work.
Unsolicited Direct Confession	A subordinate naval officer had missed a training session, when asked the reason he lied. Several hours later the subordinate located the respondent, confessed, and appologized for the deception.
Verbal/Nonverbal Behavior	NO RESPONSES
Inconsistencies with Knowledge	A subordinate was supposed to complete a task and stated they had however, the respondent had previously checked the status and knew it had not been completed.
Combination	A subordinate lied about drug use which was proven true by a positve urinalysis, subsequent eye-witness statements, and a drop in the individual's performance.
Other	An enlisted soldier was caught with drugs at the airport, the resulting court martial sentence was basis for discovery of the lie after the fact.

The second portion of the questionnaire regarded the amount of time, in months, that had passed since the respondent had been told the lie. The range provided was 0-96 months with the mean being 14.7228 and a standard deviation of 22.19018 as shown in Table 8.

Table 8. "How Long Ago" Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
How Long Ago	83	0	96	14.7228	22.19018
Valid N	83				

Analysis was then conducted on the relationship of the participant and the liar as displayed in Table 9. Respondent disclosure analysis shows the primary relationship occurred between participants and subordinates (38.1%) with the least reported relationship between participants and superiors/immediate bosses (8.3%).

Table 9. Frequencies of Participant and Liar Relationship

	ƒ	%
None Specified	36	42.9
Superior/Immediate Boss	7	8.3%
Subordinate	32	38.1%
Child	1	1.2%
Peer/Friend	8	9.5%
Total	84	100%

The final question related to the relationship between the respondent being told the lie and the subsequent realization and discovery the lie had been told. In 69 (83.1%) of the cases, participants either did not answer the question or the answer provided was too vague for objective discernment. Of the lies told, 8 (9.6%) were the product of immediate veracity judgments. Of note, the possibility of bias arises in the qualitative results as the participants were cued toward applicable near-term, work-related lies. The remainder of the timing data can be found in Table 10.

Table 10. Frequencies of Lie Time Lapse

	ƒ	%
No Answer Provided	69	83.1
Immediate Detection	8	9.6%
< One Day	3	3.6%
< One Week	3	3.6%
Total	83	100.0%

In the highly charged culture of the military, where lies carry significant consequences, it is expected there will be a high participant and subordinate discovery method response. Park et al. (2002) supports this with "…it can be argued that accuracy should be higher for high stakes lies because there should be more nonverbal leakage when the stakes are high" (p. 146). This is increased if consequences are shared by the respondent. Examinations of further distinguishing discovery methods are detailed in Chapter IX.

THIS PAGE INTENTIONALLY LEFT BLANK

IX. STUDY 2: DISCUSSION, FINDINGS, LIMITATIONS, AND RECOMMENDATIONS BASED ON ANALYSIS

What is notable about this study's findings in addition to Park et al.'s (2002) is that within the domestic military officer population there were no reported instances of at-the-time verbal/nonverbal discovery methods used. This is in direct contradiction of the prevailing historically held concept as outlined in the literature review, that deception detection primarily coincides with the lie being told and is based on the verbal and nonverbal cues elicited by the liar.

A. DISCUSSION AND FINDINGS

As hypothesized, this discovery method results closely resembled those of Park et al. (2002), 3rd party and physical information were the most prevalent discovery methods in both our works. Table 11 is provided for comparison.

Table 11. Frequencies of Recalled Lie Discovery Methods (from Park et al., 2002)

Discovery Method	ƒ	%
None Listed	-	-
Third Party Information	62	32.0%
Physical Information	35	18.0%
Solicited Direct Confession	7	3.6%
Unsolicited Direct Confession	16	8.2%
Verbal/Nonverbal Behavior	4	2.1%
Inconsistencies with Knowledge	4	2.1%
Inadvertent Confession	4	2.1%
Combination	60	30.9%
Other	2	1.0%
Total	194	100.0%

B. LIMITATIONS

As this study was directed at occupational deception, the frequency of subordinates as a participant in deception was a significant and expected departure from Park et al's primacy of social deception. This is also a possible limitation as other, more

recent, higher stakes deception outside of the workplace may have occurred but was not reported and therefore not analyzed.

Although Park et al.'s (2002) method was replicated, the resulting research indicates prevalence in the sample of immediate detection whereas Park et al.'s (2002) findings show "few recalled lies were detected on the basis of at-the-time source behaviors and that the deception detection time line is often much longer than is assumed commonly" (p. 156). The resulting disparity may be due to the current study's focus on workplace deception where the stakes are nominally greater and deception more costly.

The reader is cautioned regarding these findings as in each of the research question results a large percentage of the population did not provide responses. This could be a factor of participant fatigue as the experiment combined quantitative (Study 1) and qualitative (Study 2) aspects in one sitting.

C. RECOMMENDATIONS BASED ON ANALYSIS

The use of third party, physical, and combinatorial detection methods illustrates the applicability of the approach when the detector is privy to significant amounts of pertinent information. Access to such information provided for timely realization of deception, as noted previously in Table 10. The greater the amount of diagnostically useful information, the more rapidly the sample was able to gauge veracity. This illustrates the fallacy of verbal and nonverbal cues while reinforcing the importance of historical source material in truth-lie determination. Therefore, follow-on and remedial training for both expert and inexpert detectors should include a focus on pertinent information collection methods and associated use in the making of veracity judgments. The next chapter discusses the summary, conclusions and areas for further research.

X. SUMMARY, CONCLUSIONS, AREAS FOR FURTHER RESEARCH

The successful replication of previous studies using the unique sample of domestic military officers supports the argument of diagnostic utility usage over cue-based deception detection. With respect to Park et al's (2002) findings "Despite both the sheer quantity of research and the consistency of findings there is a history of researchers voicing dissatisfaction with various features of earlier studies," the authors' goal to improve this consistency through both quantitative and qualitative reasoning has been satisfactorily accomplished (p. 144).

A. SUMMARY

The purpose of this research was to conduct the experimental replication and documentation of further improvement in deception detection accuracy over that of previous findings. Examining the applicability of the Levine et al. (2014) methodology of content and contextual based questioning strategies in deception detection coupled with Park et al.'s (2002) approach of defining what information military officers utilized when judging veracity can effectively support this experimental replication. The nature of this topic dictates the use of both qualitative and quantitative opportunities to further expand known knowledge of deception detection. Quantitatively, the focus was on strategic questioning methods and resulting deception detection accuracy. Qualitatively, this research looked at when, how, and with what information deception is actually detected in the workplace. The following sections provide answers to the research questions in this study based on the analysis in chapter V for Study 1 and in chapter IX for Study 2.

Utilizing the diagnostic utility methods of content and contextual questioning presented in the Levine et al. (2014) study, are Department of Defense officers able to distinguish deception with greater accuracy than the previous meta-analysis mean of 54% and if so to what degree?

According to the analysis of the data culled from Study 1, the results of Levine et al.'s (2014) study of diagnostic utility and judge expertise are highly replicable with U.S. military officers. The importance of this is that the findings exhibited an increase in deception detection irrespective of expertise. Both expert and inexpert categories saw a marked increase in accuracy. In light of the literature review, typical physical and nonverbal variables saw a replicable meta-data analysis mean of 54% (Bond & DePaulo, 2006), while the authors' replicable method of manipulating question effects through the use of applied content and contextual questioning raised expert participant accuracy to 75.7% and inexpert participant accuracy to 81.0%. The accuracy increase of this replication using the previously defined sample of domestic military officers actually saw a notable increase of almost 10% in accuracy for inexpert participants over the Levine et al. (2014) results.

Utilizing the Park et al. (2002) method, what types of information do U.S. military officers report using when detecting lies in in the workplace?

Analyzing the data from Study 2, it was apparent the sample of domestic military officers used 3rd party and physical information as the primary independent methods for detecting lies in the workplace. The predominance of lie recognition occurred within immediately and involved relationships between participants and workplace subordinates. These findings help ascertain the nature of honesty in the workplace and reinforce the need for domestic service members to be properly trained in proper deception and fraud detection techniques regardless of their primary duties.

B. CONCLUSIONS

The quantitative and qualitative studies conducted during the course of this thesis have broadened the scope of deception detection with regards to domestic military officers and their determination of veracity judgments made in the workplace. Quantitatively, in the ability to replicate the Levine et al. (2014) findings in study 1, the results exhibit a positive applicability of questioning effects, namely diagnostic utility, when crafting deception detection activities. The significance of the findings combined with a lack of a main effect, that expertise is not a factor, proves the applicability of

Study 1 to a wide range of customers. Knowing that sophisticated training does not lead to a deception detection advantage is a sound financially feasible benefit for future use of this research. On a qualitative level, Study 2 results suggest and reinforce the use of specific types of information that is required to actually detect deception in the workplace with the emphasis placed on individual detectors' specific backgrounds and historical subjectivity. Thus, in order to achieve the desired accuracy, additional information is needed for the majority of the selected population to correctly deduce deception realization. Levine et al.'s (2014) comments support the value of research replication fundamentally with: "Thus, we see no reason to believe that our concept is limited to the cheating experiment context" (p. 286). The culmination of this thesis is the agreement of the authors' with that of Levine and the subsequent expansion of the research to that of DOD applicability.

C. AREAS FOR FURTHER RESEARCH

The single context and limited population of the two conducted studies presents inherent limitations but in doing so provides opportunity for further research. The next step is to refine the questioning for specific customers. Such entities' duties should include some form of high-stakes interrogation such as the military intelligence community regarding defector/counterintelligence questioning and financial examiners investigating high-dollar fraud. Follow-on research should seek to examine if this same effect is replicable in the international officer population at NPS. In doing so, such factors as region, predominant language, and cultural norms should be observed and included as demographic variables of interest. An in-depth international sample would also lend itself to the tailoring of strategic questioning methods used to investigate defector veracity. Though the purpose of this study was to assess the replicability and applicability of the diagnostic utility method of questioning, further refinement of scenarios specifically tailored to military and fraud examination should be crafted. Though military counterintelligence techniques will not be discussed due to classification issues, the reliance of the Association of Fraud Examiners fraud examination curriculum (Association of Certified Fraud Examiners, 2014) on the outdated methods of nonverbal variables for deception detection could benefit greatly from the inclusion of the

diagnostic utility method. Further research should examine the possibility of improving Study 2 participation by conducting the survey separately.

Military intelligence, counterintelligence, and defector questioning

A primary step in evaluating use in military intelligence and defector questioning is to examine wither the methods and results used in this research are applicable to international military officers and to what degree, positive or negative. Further enhancement and subsequent analysis should be made in order to ascertain the effect of, and incorporate, improved psych-cultural factors which account for differences in language proficiency, education, and biases towards deception as societally defined. This analysis should be conducted on a regionally aligned level and then further refined for individual countries on the basis of statistical significance. Additionally, the Levine et al. (2014) finding of "negative utility answers hindering experienced judges more than positive utility answers assisting experience judges" which holds true in this experiment, should be further analyzed using a robust sample of deception detection experts. While this research will be relatively straight forward, the real significance will be the evaluation of non-Western cultures and societies. In furtherance of these studies and the broadening of the deception detection field, future researchers are guided to Alexander L. George's (1967) work in analyzing Chinese Communist Army defectors during and after the Korean War; Leon Goure's (1968) RAND and Advanced Research Projects Agency (ARPA) sponsored research into Viet Cong and North Vietnamese defector questioning; Granhag and Stromwall's (2001) work on interpersonal deception as applied to interrogations; and lastly Levine et al.'s (n.d.) White Paper for the U.S. Army's Intelligence Center and School on the military applicability of diagnostic utility for use in investigation and interrogation.

Department of Defense and Civilian Fraud Investigation

From 2001 through 2011, the Undersecretary of Defense for Acquisition, Technology, and Logistics, via the Department of Justice, identified 54 DOD contractor companies charged with criminal fraud and 300 DOD affiliated companies who had civil judgments for fraud rendered against them which in total defrauded the DOD of $1.1

trillion (Department of Defense, 2011). Fraud on this scale is nothing short of shameful, what is most incriminatory, however, is the duration over which these acts took place. Without effective internal controls and appropriately trained government acquisition personnel in fraud deterrence and detection, the vulnerability for fraud will continue to exist (Rendon & Rendon, in press). Correspondingly, the Association of Certified Fraud Examiners' *2014 report to the nations on Occupational Fraud and Abuse* studied 1,483 cases of civilian fraud resulting in "a potential projected global fraud loss of nearly $3.7 trillion" (p. 4). Regardless of the public or private arena, further research should be conducted using DOD acquisition professionals and certified fraud examiners in order to ascertain the abilities of the population and the applicability of diagnostic utility in their investigations to stem the tide of fraudulent deception. An additional study, modeled on our Study 2, should be applied to determine what factors this population uses to make veracity judgments.

APPENDIX. CONSENT FORM AND QUESTIONNAIRE

Consent Form

You are invited to participate in a research study to measure your perceptions about others' communication. The purpose of the research is to better understand how people's perceptions of communication impact their judgments. Your participation should take about 30 minutes to complete.

You will be asked to watch short video clips, make judgments about each clip, and complete a survey about past experiences you have had with similar communication situations.

Your participation is voluntary. If you participate, you are free to skip any questions or stop participating at any time without penalty. The alternative to participating in the research is to not participate.

Your responses are anonymous and will not be linked to your identity in any way. No personally-identifying information will be collected – the survey only asks for broad demographic information and no other identifiers from participants.

The anticipated benefit from this study is that the findings will contribute to a larger body of knowledge, and will be used to inform coursework at NPS. You will not directly benefit from your participation in this research.

There are no known or anticipated risks associated with participation.

Results of the survey will be used responsibly and protected against release to unauthorized persons; however, there is a minor risk that data collected could be mismanaged. Only the researchers will have access to the data which will be stored on a password-protected computer.

If you have questions regarding the research, or if you experience any injury or discomfort, contact Dr. Lisa Lindsey, LLindsey@nps.edu, Principal-Investigator. If you have any questions regarding your rights as a research subject, please contact the Naval Postgraduate School IRB Chair, Dr. Larry Shattuck, 831.656.2473, lgshattu@nps.edu.

Statement of Consent. I have read the information provided above. I have been given the opportunity to ask questions and all the questions have been answered to my satisfaction. I agree to participate in this study. I understand that by agreeing to participate in this research and checking the box below, I do not waive any of my legal rights.

☐ I consent to participate in the research study.

☐ I do not consent to participate in the research study.

You will see three sets of videotaped interviews. The basic situation is always the same, but the interviewer, the person interviewed, and the questions are different.

Background: These clips are of interviews with college students who participated in a study about teamwork. Each subject had just played a trivia game with a partner for a cash prize. All participants were given an opportunity to cheat when the experimenter was called out of the room, and the answers were left in a folder within easy reach of the participants. Some participants cheated and others did not. All the people being interviewed on these tapes denied cheating.

Instructions: Watch each interview and decided if you think they cheated or not. For each interview, circle an answer indicating your opinion about whether you think that they were honest and didn't cheat or that they really did cheat and are lying about not cheating.

Set 1 Video: Exline1_4clips (6 min.)

Number	Interview	Judgment (circle one)	
1	39	Honest non-cheater	Cheated and lying about it
2	45	Honest non-cheater	Cheated and lying about it
3	44	Honest non-cheater	Cheated and lying about it
4	54	Honest non-cheater	Cheated and lying about it

Set 2 Video: Exline2_4clips (11 min)

Number	Interview	Judgment (circle one)	
5	54	Honest non-cheater	Cheated and lying about it
6	57	Honest non-cheater	Cheated and lying about it
7	71	Honest non-cheater	Cheated and lying about it
8	72	Honest non-cheater	Cheated and lying about it

Set 3 Video: Exline4_4clips (12 min)

Number	Interview	Judgment (circle one)	
9	25	Honest non-cheater	Cheated and lying about it
10	18	Honest non-cheater	Cheated and lying about it
11	10	Honest non-cheater	Cheated and lying about it
12	12	Honest non-cheater	Cheated and lying about it

*Finally, we'd like for you to recall a recent **work-related** situation in which you discovered that someone lied to you. Please take a moment to think of an example and remember as much as you can about what happened. Keeping this situation in mind, please answer the following questions:*

1. Recall as much as you can about the situation in which the person originally lied to you. In as much detail as possible, describe the event where you were lied to: where did it happen? What was the lie about? If you can, be sure to write down the exact thing that the person said to you.

2. How long ago did this event (the lie) originally take place?

3. What was/is the relationship between you and the person who lied to you?

4. Now, think about how you found out that the person lied to you. Describe in as much detail as you can the events surrounding your discovery of the lie: how exactly did you find out that the person lied to you?

Please tell us about yourself (circle the correct answer or fill in the blank):
Sex: Male Female Age: _____ Years of Military Service: _____

Rank:_____ Branch (circle one): Army Navy Air Force Marines

Have you ever received formal interviewing or interrogation training? No Yes

Have you ever conducted interviewing or interrogation as a regular part of your job? No Yes

Are you: U.S. Military International Military (please specify country) _____

Race/Ethnicity: Caucasian/White
African American/Black
Asian/Pacific Islander
Hispanic/Latino(a)
American Indian
Alaskan or Hawaiian Native
Other (please specify) _____

LIST OF REFERENCES

Association of Certified Fraud Examiners. (2014). *2014 fraud examiners manual* (US ed.). Austin, TX: ACFE Inc.

Association of Certified Fraud Examiners. (2014). *2014 Report to the Nations on Occupational Fraud and Abuse:2014 Global Fraud Study*. Austin, TX: ACFE Inc.

Bond, C. J., & DePaulo, B. M., (2006). Accuracy of deception judgments. *Personality and Social Psychology Review*, 214–234.

Buller, D. A. (1996). Interpersonal deception theory. *Communication Theory*, 203–242.

Burgoon, J. K., Buller, D. B., Ebesu, A. S., & Rockwell, P. (1994). Interpersonal deception V: Accuracy in deception detection. *Communication Monographs*, *61*, 303–325.

CNA Corporation. (2012). *Population representation in the military services fiscal year 2010* (38th Annual Report on Military Demographics). Washington, DC: U.S. Government Printing Office.

Clemens, F., Granhag, P. A., & Strömwall, L. A. (2013). Counter-interrogation strategies when anticipating questions on intentions. *J. Investig. Psych. Offender Profil.*, *10*, 125–138. doi: 10.1002/jip.1387

Department of Defense. (2011). *Report to Congress on Contracting Fraud.* (RefID: D-315854E). Retrieved from http://www.sanders.senate.gov/imo/media/doc/102011%20-%20DOD%20Fraud%20Report.pdf

DePaulo, B. M., Kirkendol, S. E., Tang, J., & O'Brien, T. P. (1988). The motivation impairment effect in communications of deception: Replications and extensions. *Journal of Nonverbal Behavior, 12*, 177–202.

Feeley, T. H., deTurck, M. A., & Young, M. J. (1995). Baseline familiarity in lie detection. *Communication Research Reports*, *12*, 160–169.

George, A. L. (1967). *The Chinese Communist Army in Action: The Korean War and Its Aftermath*. New York, NY: Columbia University Press.

Goure, L. (1968). Inducements and Deterrents to Defection: An Analysis of the Motives of 125 Defectors. (RM-5522-1-ISA/ARPA). Santa Monica, CA. The RAND Corporation.

Granhag, P. A., & Stromwall, L. A. (2001). Deception Detection: Interrogators' and Observers' Decoding of Consecutive Statements. *The Journal of Psychology*, 135 (6), 603-620.

Park, H. S., Levine, T., McCornack, S., Morrison, K., & Ferrara, M. (2002). How people really detect lies. *Communication Monographs*, *69*(2).

Kalbfleisch, P. J. (1994). The language of detecting deceit. *Journal of Language and Social Psychology*, *13*, 469–496.

Kraut, R. (1980). Humans as lie detectors. *Journal of Communications*, *30*, 209–216.

Levine, T. R., Blair, P. J., & Clare, D. D. (2014). Diagnostic utility: Experimental demonstrations and replications of powerful question effects in high-stakes deception detection. *Human Communication Research*, 262–289.

Levine, T. R., Lindsey, L. L. M., Blair, P. J. (n.d.). *An Experimental Examination of the Effects of Expert Interrogation on Deception Detection and Confession Rates*. White Paper prepared for the United States Army Intelligence Center & School.

Lindsey, L. L. M., Dunbar, N. E., & Russell, J. C. (2011). Risky business or managed event? Perceptions of power and deception in the workplace. *Journal of Organizational Culture, Communication and Conflict*, *15*(1), 55–79.

Millar, M., & Millar, K. (1995). Detection of deception in familiar and unfamiliar persons: The effects of information restriction. *Journal of Nonverbal Behavior*, *19*, 69–84.

Miller, G. R., & Stiff, J. B. (1993). *Deceptive communication*. Newbury Park, CA: Sage.

Rendon, R.G. and Rendon, J.M. (in press). Auditability in public procurement: An analysis of internal controls and fraud vulnerability, Int. J. Procurement Management

Stiff, J. B., & Miller, G. R. (1986). "Come to think of it…": Interrogative probes deceptive communication and deception detection. *Human Communication Research, 12*, 339–357.

Vrij, A. (1994). The impact of information and setting on detection of deception by police detectives. *Journal of Nonverbal Behavior, 18*, 117–136.

Vrij, A. (2000). *Detecting lies and deceit: The psychology of lying and the implications for professional practice*. West Sussex, UK: John Wiley and Sons.

Trovillo, P. V. (1939). History of Lie Detection. *Journal of Criminal Law and Criminology*, 848–881.

INITIAL DISTRIBUTION LIST

1. Defense Technical Information Center
   Ft. Belvoir, Virginia

2. Dudley Knox Library
   Naval Postgraduate School
   Monterey, California